

UNITED STATES PATENT APPLICATION

for

**WAVELET-BASED IMAGE PROCESSING PATH**

Inventors:

**Kathrin Berkner**  
**Edward L. Schwartz**  
**Michael Gormish**

prepared by:

**BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP**  
12400 Wilshire Boulevard  
Los Angeles, CA 90025-1026  
(408) 720-8598

File No.: 74451.P110

**EXPRESS MAIL CERTIFICATE OF MAILING**

"Express Mail" mailing label number: EL143566688US

Date of Deposit: September 8, 2000

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Assistant Commissioner for Patents, Washington, D. C. 20231

Mara E. Brown

(Typed or printed name of person mailing paper or fee)

Mara E. Brown  
(Signature of person mailing paper or fee)

9/8/00  
(Date signed)

## WAVELET-BASED IMAGE PROCESSING PATH

### FIELD OF THE INVENTION

The present invention relates to the field of image processing; more 5 particularly, the present invention relates to a wavelet-based image enhancement system.

### BACKGROUND OF THE INVENTION

Image processing has many applications. One such application is in a 10 digital copier. A digital copier typically has different modes for copying text or an image. The different modes, text or letter mode vs. photo mode, include special image processing techniques such as sharpening or smoothing of the scanned document, thinning of text, despeckeling, etc. Problems arise if the copied document has text and image content. In order 15 to apply different techniques to text and image regions of a document, a segmentation of the document or a classification of the pixel of the scanned document is needed. Without out this classification, an oversmoothing of text in the photo-mode may occur and an enhancement of halftone noise in the text mode may occur. A consequence of this enhancement of halftone 20 noise in the text mode can be the creation of severe Moire patterns after resolution reduction. Such artifacts are observed in many current digital

copiers. A typical imaging path for digital gray scale copier is demonstrated in Figure 1.

Referring to Figure 1, a scanned image 100 is input into processing block 101, which changes gray values in the scanned image to compensate for scanner bias through lookup tables or other point operation (e.g., gamma correction). Next, different regions in the scanned image are classified (e.g., text, halftone, continuous tone, etc.) (processing block 102). The pixels are then filtered by processing block 103 depending on the classified pixel. Such filtering may include smoothing or sharpening of pixels based on their classification. After filtering, the pixels undergo up and/or downsampling by processing block 104. The pixels may also be subjected to resizing for enlarging/reducing. Next, the gray values of the pixels are changed through look up tables or other point operation (e.g., gamma correction) to compensate for printer bias (processing block 105). Once changing of the gray values has been done, halftoning is performed on the pixels (processing block 106). For non-copier applications, halftoning might be replaced with another output formatting step, the result of which is processed image 107.

Single components of the digital imaging copier path are prior art, such as contrast enhancement of the scanned image or halftoning for

printing. See U.S. Patent No. 5,883,973, entitled "Method and Apparatus for Processing a Document by Segmentation into Text and Image Areas," issued March 16, 1999 to Pasovici and Shu, and U.S. Patent No. 5,805,721, entitled "Method and Apparatus for Contrast Enhancement," issued September 8, 5 1998 to Vuylsteke and Schoeters. There are also publications on image enhancement with wavelets that include denoising and contrast enhancement. For example, Zong, Laine, Geiser, and Wilson, "De-Noising and Contrast Enhancement via Wavelet Shrinkage and Non-linear Adaptive Gain," Proceedings of the SPIE, Vol. 2762, pgs. 566-574, 1996. Image 10 enhancement using the Laplacian pyramid to add high frequency content to the image is also known. For example, see U.S. Patent No. 5,717,789, entitled "Image Enhancement by Non-linear Extrapolation in Frequency Space," issued February 1998 to Anderson and Greenspan.

**SUMMARY OF THE INVENTION**

A system is described. In one embodiment, the system comprises an image processing path and a print engine coupled to the image processing path. The image processing path processes an input image in a wavelet 5 domain.

00000000000000000000000000000000

**BRIEF DESCRIPTION OF THE DRAWINGS**

The present invention will be understood more fully from the detailed description given below and from the accompanying drawings of various embodiments of the invention, which, however, should not be taken 5 to limit the invention to the specific embodiments, but are for explanation and understanding only.

**Figure 1** illustrates a prior art image processing path for a grayscale copier.

10

**Figure 2A** is a block diagram of one embodiment of a system having a wavelet-based image processing path.

15

**Figure 2B** is a block diagram of one embodiment of a digital copier.

**Figure 2C** is a block diagram of one embodiment of a wavelet-based processing path for a digital copier.

20

**Figure 3** illustrates a classifier for use in a copier or printer.

**Figure 4A** illustrates the formation of wavelet coefficients of a halftone and diagonal line pattern using a 2D Haar wavelet transform.

**Figure 4B** illustrates a schematic overview of the classification based on the first level wavelet coefficients.

**Figure 5** illustrates 3x3 windows in the LH, HL, HH bands at level 5 one of the wavelet decomposition.

**Figure 6** illustrates exemplary coefficient templates.

**Figure 7** is a block diagram of one embodiment of a copier.

10

**Figure 8** is a block diagram of one embodiment of a computer system.

## DETAILED DESCRIPTION OF THE PRESENT INVENTION

A method for processing an image and apparatus for doing the same are described. In the following description, numerous details are set forth in order to provide a thorough understanding of the invention. It will be

5 apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

Some portions of the detailed descriptions which follow are presented

10 in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-

15 consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for

reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are

5 merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a computer

10 system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

15 The present invention also relates to apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable

storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic

5 instructions, and each coupled to a computer system bus.

The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized

10 apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the

15 invention as described herein.

A machine-readable medium includes any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory ("ROM"); random access memory ("RAM"); magnetic disk storage media;

optical storage media; flash memory devices; electrical, optical, acoustical or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.); etc.

## 5 Overview

The processing described herein provides for image processing between scanned (or otherwise sensed) digital images and digital images for display or printing. This image processing might be in a digital image in, digital image out device such as, for example, a printer controller (e.g., for a network) or performed in a general purpose computer. The image processing described herein might be combined with an analog input device in, for example, a scanner or digital camera, with a print engine in a printer or with another output device, or with both an input and output device, for example in a digital copier, facsimile machine or multifunction machine.

The wavelet-based image processing described herein may be used to compensate for differences between input and output systems. Some image in and image out systems have matched input and output devices that require little or no image processing between input and output for particular types of images. For example, some video cameras and video displays are

well matched and good results for images of natural scenes require little or no intermediate processing. In contrast, some systems have dissimilar input and output devices or complex/compound image types where significant processing is required. For example, an input device might require less

5 noise, higher sharpness and/or a different size/scale than the input device provides. Additionally different images or different positions in the same image might have different enhancement. The wavelet-based image processing described herein may be used to process an image from an input system (e.g., device) so that it "matches" the output system (e.g., device).

10

#### Overview of a Wavelet-based Image Processing System

A wavelet-based imaging path for a digital copier or multifunction machine having automatic adaption of image processing routines to various characteristics in the image such as, for example, text and background is

15 described.

The wavelet-based image processing may occur after non-wavelet-based processing blocks or modules and/or before other non-wavelet-based processing blocks or modules. Figure 2A is a block diagram of a system having a wavelet-based image processing path. Referring to Figure 2A, an

image source 250 generates image data. Image source 250 may comprise, for example, a digital camera, scanner, a network or channel, storage, etc. Then processing logic, optionally, changes gray values through look up tables (e.g., gamma correction) (processing block 251). Alternatively, the image

- 5 data may be subjected to a point operation. Afterwards, processing image data is received by wavelet-based image processing block 252 for wavelet-based processing. After wavelet based processing, processing logic optionally changes the gray values of the image data through look up tables (e.g., gamma correction) or, alternatively, a point operation. Next,
- 10 processing logic performs halftoning on the image data or some other form of preparation for output (processing block 254) and delivers the processed image data to an image distinction 255. Image distinction 255 may comprise, for example, a printer, print engine, display, network or channel, or storage.

Figure 2B is a block diagram of a copier. Referring to Figure 2B,

- 15 processing logic receives a scanned image 280, and then, optionally, changes gray values in the scanned image 200 through look up tables (e.g., gamma correction) (processing block 281). Alternatively, scanned image 280 may be subjected to a point operation. Afterwards, processed image data is received by wavelet-based image processing block 252 for wavelet-based processing.

After wavelet-based processing, processing logic optionally changes the gray value of the image data through look up tables (e.g., gamma correction) or, alternatively, a point operation. Next processing logic performs halftoning on the image data or some other form of preparation for output

5 (processing block 254), thereafter outputting processed image 285.

Figure 2C is a block diagram of one embodiment of a wavelet-based image processing path for a digital copier. In one embodiment, this image processing path replaces processing blocks 102-106 in Figure 1 in a copier or other image processing device. The processing may be performed by

10 processing logic that may comprise hardware, software or a combination of both.

Referring to Figure 2C, forward wavelet transforms block 201 performs a wavelet transform on the scanned image to create a set of coefficients. The coefficients may be optionally sent to wavelet-based

15 classifier processing block 202A, which classifies the coefficients.

Alternatively, image data may be directly input to a non-wavelet-based classifier block 202B. The classifier information output from classifier processing blocks 202A and 202B may be used in cooperation with

processing block 203-206, 209, 210 and 202C. These processing blocks are described below.

Processing logic performs thresholding on the coefficients output from wavelet transform block 201 (processing block 203). After

5 thresholding, processing logic rescales coefficients (processing block 204).

Next, processing logic filters coefficients (processing block 205).

After filtering coefficients, processing logic may optionally perform upsampling or downsampling (processing block 206). Subsequently, processing logic performs an inverse wavelet transform on the coefficients

10 (processing block 207). In an alternative embodiment, processing logic performs an inverse wavelet transform that includes upsampling and/or downsampling (processing block 208). Note that if processing logic performs an inverse wavelet transform in combination with up/downsampling (i.e., processing block 208), the processing block 206 is

15 not performed.

After performing an inverse transform, processing logic changes gray values on the image data through a look up table or other point operation (e.g., gamma correction) (processing block 209). Processing logic then performs halftoning on the image data (processing block 210). The output of

halftoning block 210 is a processed image 220. Note that processing blocks 209 and 210 are not performed in the wavelet domain.

The use of signal-adaptive nature of wavelet domain image processing, including a description of each of these blocks follows.

5

## Critically Sampled or Overcomplete Discrete Wavelet Decomposition

Initially, in the image processing path, forward wavelet transform block 201 performs a wavelet decomposition on the scanned image up to a chosen level  $L$ .  $L$  may be any number of levels, but 2 or 3 levels is sufficient for many applications. This transform may be critically sampled or overcomplete. Besides overcomplete (fully redundant) and critically sampled wavelet transforms, a transform with reduced redundancy or a complex transform could also be used. An example for reduced redundancy is the computation of all phases of the critically sampled wavelet transform at the first level of resolution and only some of the phases, such as every other phase at the second level of resolution (see Selesnik, I. W. "Smooth Wavelet tight Frames with Limited Redundancy," Technical report, Dept. Electrical Engineering, Polytechnic University, Brooklyn, NY, August 1999). Alternatively, the phases could be chosen dynamically, e.g., by selecting

those that contain the most energy. For example, see Coifman, R. R.,  
Donoho, D. L. "Translation-invariant De-noising," in *Wavelets and Statistics*  
(Editor A. Antoniades), Springer Lecture Notes, Springer-Verlag, 1995.

Complex transforms compute real and imaginary parts wavelet coefficients,

5 which results in a redundancy factor of two. The imaginary parts contain  
phase information on an image region. For example, see Kingsbury, N. "The  
dual-tree complex wavelet transform: a new efficient tool for image res-  
toration and enhancement," Proceedings European Signal Processing Conf.,  
1998, pp. 319-322.

10 If the image is given by NxN data points then, when using a critically  
sampled wavelet transform for L levels, the decomposition is also given by  
NxN data points, namely wavelet and scaling coefficients. When using the  
overcomplete wavelet transform for L levels, the decomposition is given by  
(3L+1)NxN data points.

15 In both cases, wavelet coefficients  $d_l$  for levels  $l=1$  to L and scaling  
coefficients  $c_l$  at the coarsest level have to be stored. In one embodiment, an  
overcomplete Haar wavelet transform is used. An example of a 1D critically  
sampled Haar transform (without  $L_2$ - normalization, low pass =  $a + b$ , high  
pass =  $a - b$ ; ignore boundaries) is given below:

Table 1

input	low pass output	high pass output
5	12	-2
7		
1	2	0
1		
8	15	1
7		

An example of overcomplete Haar transform is as follows:

Table 2

input	low pass output	high pass output
5	12	-2
7	8	6
1	2	0
1	9	-7
8	15	1
7		

Examples of filter coefficients for different wavelet systems are as follows:

Haar:  $[1 \ 1][1 \ -1]$  normalized by  $1/\sqrt{2}$ .

2-6:  $[1 \ 1][-1/8 \ -1/8 \ 1 \ -1 \ 1/8 \ 1/8]$  normalized by  $1/\sqrt{2}$ .

5-3:  $[-1/4 \ 1/2 \ 3/2 \ 1/2 \ -1/4][1/4 \ -1/2 \ 1/4]$  normalized by  $1/\sqrt{2}$ .

2-10:  $[1 \ 1][3 \ 3 \ -22 \ -22 \ 128 \ -128 \ 22 \ 22 \ -3 \ -3]/128$  normalized by  $1/\sqrt{2}$ .

9-7:  $[0.037828 \ -0.023849 \ -0.110624 \ 0.377402 \ 0.852699 \ 0.377402$   
 $-0.110624 \ -0.023849 \ 0.037828], [0.064539 \ -0.040689 \ -0.418092$   
 $0.788486 \ -0.418092 \ -0.040689 \ 0.064539]$

daub6:  $[0.0352 \ -0.0854 \ -0.1350 \ 0.4599 \ 0.8069 \ 0.3327], [-0.3327 \ 0.8069$   
 $-0.4599 \ -0.1350 \ 0.0854 \ 0.0352]$ .

The forward low pass and high pass filters are specified. For

symmetric biorthogonal filters, the inverse filters are the forward filters with

5 alternating signs changed starting with the second and reversed. For example for the 2-6 filter:

Forward low pass  $(1 \ 1) \frac{1}{\sqrt{2}}$

Forward high pass  $(-1/8 \ -1/8 \ 1 \ -1 \ 1/8 \ 1/8) \frac{1}{\sqrt{2}}$

Inverse low pass  $(-1/8 \ 1/8 \ 1 \ 1 \ 1/8 \ -1/8) \frac{1}{\sqrt{2}}$

10 Inverse high pass  $(-1 \ 1) \frac{1}{\sqrt{2}}$

This selection of filters covers the characteristics for the most widely used wavelet families (orthogonal, biorthogonal, symmetric, even-length,

odd-length). The transform may also be any LeGaul-Tabatabai biorthogonal transform.

For a two dimensional wavelet transform, both low pass and high pass filters are applied horizontally and vertically. The four resulting

5 subbands are:

	<u>Subband</u>	<u>horizontal filter</u>	<u>vertical filter</u>
10	LL	low pass	low pass
	HL	high pass	low pass
	LH	low pass	high pass
	HH	high pass	high pass

### Multiscale Sharpening, Smoothing and Denoising

In one embodiment, thresholding and multiplication are performed  
 15 on the wavelet coefficients to achieve denoising, sharpening, and smoothing. A criterion dictates whether thresholding or multiplication are applied. In one embodiment, the criterion comprises the magnitude of wavelet coefficients. Alternatively, a classifier could be used to indicate whether to apply thresholding and/or multiplication.

20

#### *Denoising by Thresholding*

Denoising may be performed by setting all coefficients below a given threshold to zero. This technique is known as hard-thresholding and is well-

known in the art. As an alternative, soft-thresholding, e.g., shrinking the coefficients above the threshold toward zeros by the amount of the threshold can also be considered. Such a technique may be advantageously used for critically sampled and overcomplete wavelet decompositions, and is

- 5 described in Donoho, "Denoising by Soft-Thresholding", IEEE Transactions on Information Theory, 41(3): 613-627, 1995. The threshold may be one that is calculated for the presence of additive Gaussian White noise in an image. This threshold depends only on the size of the data and the variance of the noise which is estimated from the wavelet coefficients using a well-known
- 10 procedure (e.g., statistical estimation of variance). Other choices of threshold selection may be used. These may be calculated from a series of test images.

There are additional strategies for thresholding available. There may be a global threshold or different thresholds for different levels of

- 15 decomposition or different bands (LH, HL, HH). The thresholds may be set manually. Thresholds may depend on one or more of the following parameters:

1. Standard deviation or median of absolute values of coefficients in one region (e.g., HH coefficients at level 1) (global threshold)

2. Standard deviation or median of absolute values of HH coefficients at each level (level-dependent threshold)

3. Standard deviation or median of absolute values of LH/HL/HH coefficients at first level (band-dependent threshold). For the example of two levels of decomposition three thresholds are chosen

5

4. Standard deviation or median of absolute values of LH/HL/HH coefficients at each level (level-band-dependent threshold). For the example of two levels of decomposition, six thresholds are chosen.

5. Manually set thresholds (global, level-dependent, level-band dependent) using controls that are part of the user interface of the device.

10

6. Local thresholding by computing local variances or medians of coefficients in between a band.

7. Given classifier: select for each strategy (global, level-dependent, level-band dependent) a number of thresholds corresponding to the number of classes. Then use those classifier-dependent thresholds.

15

Laplacian based techniques applied to wavelets do not perform the superior wavelet-only denoising. It is important to note that wavelet denoising techniques do not apply to Laplacian pyramid decompositions.

Laplacian pyramid decompositions can be used only for systems where inferior or no denoising is acceptable.

### Rescaling of Coefficients

5 After any denoising is completed, the wavelet coefficients that are non-zero are modified for the purpose of sharpening or smoothing. In one embodiment, sharpening or smoothing of the image is obtained by multiplication of wavelet coefficients with a level-dependent parameter  $\mu_j > 0$ . For sharpening, the high frequency bands should be multiplied with a

10 larger factor than the low frequency bands, whereas for smoothing the low frequency bands should be multiplied with larger factors than the high frequency bands. One way to obtain this rescaling is to multiply the coefficients at level  $j$  with  $\mu_j = R \cdot 2^{\alpha j}$ . The parameter  $R$  functions as a renormalization factor for preserving the global energy of coefficients and  $\alpha$

15 is the parameter that determines the degree of sharpening ( $\alpha < 0$ ) or smoothing ( $\alpha > 0$ ). For more information, see U.S. Patent Application Serial No. 09/467,544, entitled "Multiscale Sharpening and Smoothing With Wavelets," filed December 10, 1999, assigned to the corporate assignee and incorporated herein by reference. There are many choices for the parameter

$R$  that depend on how the energy of an image is measured. One way would

be to preserve the  $L^2$  norm  $\sqrt{\sum_{i,j} |x_{i,j}|^2}$  of the image before and after rescaling,

i. e. compute  $R$  with

$$R^2 = \left( \sum_{j,k} |d_{j,k}|^2 \right) / \left( \sum_{j,k} |\mu_j d_{j,k}|^2 \right)$$

5 If the scanned image is assumed to have a Hölder regularity of  $\gamma$  (typically  $\gamma \geq 1$ ), an approximation of  $R^2$  can be obtained by the following estimate.

$$R^2 = \left[ \frac{2^{(L+1)2\gamma} - 2^{2\gamma}}{2^{2\gamma} - 1} \right] / \left[ \frac{2^{(L+1)2(\gamma+\alpha)} - 2^{2(\gamma+\alpha)}}{2^{2(\gamma+\alpha)} - 1} \right]$$

This approximation can be computed a-priori and does not depend on the  
10 actual values of wavelet coefficients.

Instead of using the  $L^2$  norm, there are other choices for the norm.

Using the theory of Besov space, it is possible to preserve a specific Besov norm of the image. This norm is defined as a weighted sum of wavelet coefficients. The parameter  $R$  can be computed as

$$15 \quad R^q = \left( \sum_j 2^{j(\alpha+q+1/2)} \left( \sum_k |d_{j,k}|^p \right)^{q/p} \right) / \left( \sum_j 2^{j(\alpha+q+1/2)} \left( \sum_k |\mu_j d_{j,k}|^p \right)^{q/p} \right)$$

for  $0 < p, q < \infty$ .

For  $p, q = \infty$ , the term  $\left(\sum |..|^s\right)^{1/s}$  in the previous equation is substituted by

$\sup |d_{j,k}|$ .

In one embodiment, the choice of the parameters  $\alpha$  and  $R$  can be

- 5 controlled by the user interface of a device, giving the user the opportunity to select more/less sharpening, or more/less smoothing. This parameterization of sharpening and smoothing allows a fine graduation of degrees of enhancement.

As shown in Figure 2 information obtained from a classifier can be

- 10 used to determine different parameter  $\alpha_i, R_i$  for different classified areas  $A_i$  in the image.

### *Filtering of Coefficients*

The frequency content of a single subband at a specific level of

decomposition can be modified by a filtering of that subband. For example,

- 15 additional sharpening of the image could be performed by filtering the LL component at the largest level of decomposition with a sharpening filter. In a similar way, additional smoothing could be performed by applying a

smoothing filter to the LL component. Another possibility is to remove periodic patterns in a specific band by applying a special filter (e.g., a bandpass or notch filter) to that specific band/component of the wavelet decomposition. Thus, a formulation filter is applied in the wavelet domain

5 (e.g., the LL domain) that is equivalent to a larger filter in the pixel domain.

A filter may be used that is suited to reduce or remove special periodicities out of the image. The output of this filter may be used to select different denoising thresholds for halftone and text regions. Periodicity detection can be done with autocorrelation or similar functions. For

10 example,

$$f_i^j = \frac{1}{2}(|x_i - x_{i-j}| + |x_i - x_{i+j}|)$$

or

$$f_i^j = \sqrt{\frac{1}{2}((x_i - x_{i-j})^2 + (x_i - x_{i+j})^2)}$$

15 .

There is periodicity with period 2 at location i if

$$f_i^2 < f_i^1$$

There is periodicity with period 3 at location  $i$  if

$$f_i^3 < f_i^1$$

In some applications, the period (of the periodicity) is known. The

5 period might be determined from the screen size of a halftoned original to be copied and the resolution of the copier's scanner. For example, consider a period of 3. Small amplitude wavelet coefficients might be filtered in one dimension by

$$d_i^{\text{filtered}} = \frac{(d_{i-1} + d_i + d_{i+1})}{3}$$

10 For the overcomplete Haar at level 1 (and no adaptive processing such as denoising), this is equivalent to the linear filter [1 2 2 1] in the spatial domain.

A two dimensional example is

$$d_i^{\text{filtered}} = 1/9 \begin{pmatrix} d_{i-1,j-1} + d_{i,j-1} + d_{i+1,j-1} + \\ d_{i-1,j} + d_{i,j} + d_{i+1,j} + \\ d_{i-1,j+1} + d_{i,j+1} + d_{i+1,j+1} \end{pmatrix}$$

15 For the overcomplete Haar at level 1 this is equivalent to the following linear filter in the pixel domain.

$$\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 4 & 2 \\ 2 & 4 & 4 & 2 \\ 1 & 2 & 2 & 1 \end{bmatrix}$$

Similar 1x5 or 5x5 filters might be used for period 5. When even symmetric wavelet filters are used, even period periodicities are removed easily by the wavelet and extra processing is typically only desired for odd periods.

#### *Use of a Classifier and an Exemplary Classifier*

In one embodiment, denoising, sharpening and smoothing of wavelet coefficients is controlled using a classifier. A classifier could be computed 10 based on information extracted from the wavelet coefficients or from some other information source, e.g. the image pixels (Figure 2).

As a decision making step classification is part of a segmentation process that decomposes an image into various components.

Classification or segmentation can be used for many things in a 15 system. It can be used to control one or more of: de-noising, sharpening/smoothing, upsampling/downsampling, change of gray values through lookup table (e.g. gamma correction), halftoning, under-color

removal and/or image compression. Non-wavelet based filtering methods may have a single filter that performs denoising and sharpening/smoothing, and classification would select between different filters. With wavelet-based processing, not only are different parameters used for denoising and

- 5 sharpening/ smoothing, but also different scales or different wavelet subband may use different parameters. This flexibility is an advantage of wavelet based processing. Having classification only control desired parameters reduces the effect of classification errors. For example for de-noising and sharpening/smoothing in one system, perhaps only de-noising
- 10 is controlled by classification or denoising at the first transform level. In another system, perhaps denoising and sharpening/ smoothing at the first two transform levels are affected by classification.

Regardless of the source of the classification, the result can be used for a variety of things including: setting smoothing parameters, setting

- 15 sharpening parameters, choosing an "ordinary" linear filter for sharpening or smoothing, controlling the under color removal process in color copiers/printers, as an input to a document classification algorithm, as an input to a compression algorithm (like JBIG 2), as a way to control a halftoning algorithm (or to turn off halftoning for certain image types switch

between error diffusion and ordered dither). For example, when performing under color removal, one classifier may be as follows:

classify as text or line → black = min (cyan, magenta, yellow)

5 classify as not text or line → black =  $\alpha$  min (cyan, magenta, yellow) where  $0 < \alpha < 1$ , for example  $\alpha = 0.75$

In a compression framework, JBIG2 and MRC use classification in the sense that the encoder must decide how to classify, but no explicit classification scheme is specified (ISO/IEC 16485: Information technology - Mixed Raster Content (MRC)). These methods could use a wavelet-based 10 classifier in the encoder. Such a classifier is shown in Figures 3A-3C.

Classifiers can distinguish between a variety of different image area types :

- text and line art versus other
- black and white text versus colored or gray text
- 15 • text with patterned backgrounds versus text on plain backgrounds
- text of various sizes
- halftone versus other
- gathered dot halftones versus other halftones
- gathered dot halftones of various screen frequencies
- 20 • continuous-tone image versus other

Spatial domain classifiers can be mapped to the wavelet domain.

Each wavelet coefficient has a support in the spatial domain. The AND, OR or majority of the spatial domain classifications for the area of support for a coefficient can be used for the classification of a coefficient. Similarly, a

5 wavelet domain classifier can be mapped to the spatial domain. The AND, OR or majority of the wavelet domain classifications for coefficients that have support including a particular pixel can be used for the classification of a pixel.

Many kinds of classifiers can be implemented in the wavelet domain.

10 Classifiers can be based on the value of a coefficients (Figure 6A) or on statistics from regions of coefficients (Figure 6B,C,D), a more thorough explanation of which are given below. Possible statistics include:

$$\text{Mean } \bar{x} = (1/n) \sum_n x \text{ and standard deviation } \text{sd} = \sqrt{\frac{1}{n} \sum_n (x - \bar{x})^2}$$

15

- mean and first absolute moment  $\frac{1}{n} \sum_n |x - \bar{x}|$
- minimum and maximum
- minimum
- maximum

A location might be classified as text (and not halftone or continuous tone image) if the mean of region was about a threshold T1 and if the variation (standard deviation or first absolute moment) was below a second threshold T2.

5        In one embodiment, to compute a classifier from wavelet coefficients, the local standard deviations of wavelet coefficients in different bands, levels or across bands and levels, may be computed. For example, if each level has three bands (LH, HL, and HH), then the mean for a detail coefficient is:

$$ew = \frac{d_{LH} + d_{HL} + d_{HH}}{3}$$

10      The standard deviation across the bands is

$$(d_{LH} - ew)^2 + (d_{HL} - ew)^2 + (d_{HH} - ew)^2.$$

Depending on those measures, text halftone and background regions can be differentiated and a different denoising, smoothing, and sharpening parameter(s) may be applied. For example, text may be identified when the  
15      standard deviation of coefficients at one or over several levels is small, and halftone may be identified if the standard deviation is high. This makes it possible to distinguish halftone portions from lines. Thus, different denoising, sharpening or smoothing parameters may be selected for spatial

areas (e.g., local regions in wavelet components), different bands, resolutions.

In one embodiment, a halftone-vs-line classifier is used in a wavelet-based image processing path of a copier. The difference between halftoned

5 areas and text areas is that halftone areas consist of periodic patterns, whereas text consists of constant gray values in one direction.

A vertical line in the image produces large wavelet coefficients of the same sign along a vertical line in the HL band, whereas the coefficients next to the line are small. A horizontal line produces large wavelet coefficients of

10 the same sign along a horizontal line in the LH band. A diagonal line produces middle sized wavelet coefficients of the same sign along a diagonal line in the HL, LH and HH band. See Figure 4A.

On the first level of decomposition, a halftoned area produces large coefficients of positive and negative sign in the HL and LH band. These

15 coefficients are not connected, but rather "uniformly distributed". In the HH bands, a halftone area produces non-connected large size coefficients. Figure 4A illustrates the formation of wavelet coefficients of a halftone and diagonal line pattern using a 2D Haar wavelet transform.

The difference in formation of wavelet coefficients of lines and halftone areas can be captured by computing the mean and the standard deviation in a window of a special length directed in horizontal, vertical and both diagonal directions. In terms of those parameters the characterization 5 of lines and halftone areas is shown in Table 3.

Table 3 - Characterization of lines and halftone area in directional window in the wavelet domain.

	LH: mean	LH: std	HL: mean	HL: std	HH: mean	HH: std
horizontal lines	large	small	small	small	small	small
vertical lines	small	small	large	small	small	small
diagonal lines	medium	small	medium	small	medium	small
halftone	medium	large	medium	large	medium	large

10 It becomes apparent that the mean computed in a horizontal or vertical window is a sufficient criterion for separating horizontal and vertical lines from halftone or diagonal lines. For distinguishing diagonal lines from halftone, the standard deviation has to be computed as an additional criterion. This classification also works on a level larger than one, 15 if the lines have a specific thickness and the halftone a specific frequency.

However, since this is information in general not available a priory, the classification is performed only at the first level of decomposition.

In one embodiment, different classification thresholds ( $T_1, T_2, T_3, T_4$ ) are chosen to classify line and halftone areas in HL/ LH and HH bands. An 5 example for a window of length 3 is:

if (mean of horizontal window centered in LH-coefficient  $d(i,j) > T_1$  AND  
 std of horizontal window around coefficient  $d(i,j) < T_2$ )  
 then classify 3x3 window around (i,j) as being horizontal line area  
 10 else if (mean of vertical window centered in HL-coefficient  $d(i,j) > T_1$  AND  
 std of vertical window around coefficient  $d(i,j) < T_2$ )  
 then classify 3x3 window around (i,j) as being vertical line area  
 15 else if (mean of diagonal SW-NE window centered in HH-coefficient  $d(i,j) > T_3$   
 AND  
 std of diagonal SW-NE window around coefficient  $d(i,j) < T_4$ )  
 then classify 3x3 window around (i,j) as being diagonal line area  
 20 else if (mean of diagonal NW-SE window centered in HH-coefficient  $d(i,j) > T_3$   
 AND  
 std of diagonal NW-SE window around coefficient  $d(i,j) < T_4$ )  
 then classify 3x3 window around (i,j) as being diagonal line area  
 25 else  
 classify 3x3 window around (i,j) as being halftone area

A schematic overview of the classification based on the first level wavelet coefficients is shown in Figure 4B. This classification can work successfully in the line and halftone area.

Figure 5 illustrates 3x3 windows in the LH, HL, HH bands at level one of the wavelet decomposition. The bold coefficients are those used for the classifier.

In one embodiment, for construction of the classifier, the following

5 steps are followed.

1. Compute mean of the three bold values in each box in Figure 5, i.e.

$$\text{mean\_LH} = (x_4+x_5+x_6)/3,$$

$$\text{mean\_HL} = (x\_2+x\_5+x\_8)/3,$$

mean1\_HH = (x\_7+x\_5+x\_6)/3,

mean2\_HH = (x\_1+x\_5+x\_9)/3.

2. Compute standard deviation of the three bold values in each box, i.e.

std\_LH = sqrt([(x\_4-mean\_LH)^2 + (x\_5-mean\_LH)^2 + (x\_6-mean\_LH)^2]/3)

```
std_HL = sqrt([(x_2-mean_HL)^2 + (x_4-mean_HL)^2 + (x_8-mean_HL)^2]/3)
```

```
mean1_HH, ZJ/3)
std1_HH = sqrt([(x_7-mean1_HH)^2 + (x_5-mean1_HH)^2 +
```

$$(x\_3 - \text{mean1\_HH})^2] / 3)$$

std2\_HH = sqrt([(x\_1-mean2\_HH)^2 + (x\_5-mean2\_HH)^2 + (x\_9-mean2\_HH)^2]/3).

3. Set the complete  $3 \times 3$  window in the classifier image (positions of  $x_{1 \dots 9}$ ) to white

if (abs(mean LH) > thresh 1 && std LH < thresh 2) OR

if (abs(mean\_HL) > thresh\_1 && std\_HL < thresh\_2) OR

if (abs(mean1\_HH) > thresh\_3 && std1\_HH < thresh\_4) OR

if (abs(mean2\_HH) > thresh\_3 && std2\_HH < thresh\_4).

30

This version of the classifier uses the mathematical definition of standard deviation. Of course, the `sqrt` can be eliminated by adjusting the thresholds `thresh_2` and `thresh_4`. There may also be the possibility of substituting the squares  $(.)^2$  by absolute values  $||.||$ . Then the thresholds

- 5    `thresh_2` and `thresh_4` will have to be chosen appropriately.

A wavelet-based classifier could also be used for nonwavelet-based operations such as halftoning or gamma correction. One example is to detect text of particular sizes in the wavelet domain and to use OCR (optical character recognition) on the text in the pixel domain, perhaps with

- 10    parameters chosen based on the detected font size.

Wavelet coefficients from different decomposition levels can be used in a classifier.

Classifications can also effect a single coefficient or can effect a region of coefficients.

- 15    In one embodiment, a wavelet coefficient is classified as being in a text region if the statistics for it or any of its eight neighbors meet thresholds for text.

Similar to the first method, in an alternative embodiment, a region around that pixel is classified instead of a single pixel. This is essentially

looking at the effect of one pixel on a region. The size of the region may depend on the scale  $L$  (e.g., the regions are selected to be of size  $2^L \times 2^L$ ).

In yet another embodiment, for each scale  $l=1..L$ , a threshold  $T_l$  is selected. A pixel at position  $(i,j)$  is classified as text only if the coefficients at 5 the corresponding positions at all scales  $\ell=1..L$  are the corresponding  $T_\ell$  threshold. Otherwise, the pixel is classified as background.

In still another embodiment, and similar to the previous method, a complete region around the located pixel is classified as text, as opposed to 10 single pixels. Possible size of the regions include, but are not limited to:  $2^L \times 2^L$ .

In one embodiment, the classification of an image into text, half-tone, continuous tone (or other types) is performed in the wavelet domain or in the spatial domain by a variety of techniques.

## 15 Resolution Reduction and Magnification

For resampling of an image (reduction or enlargement), an interpolation filter may be used. Filters with different characteristics lead to 20 different characteristics in the resampled image. Common problems after resampling are increased moiré artifacts, less crisp text as well as broken text

and lines. One reason for the occurrence of those artifacts is that the global filter characteristics of the interpolation filter do not match the characteristics of a local region in the image. For reduction of moiré artifacts, a filter with random elements would be well suited to destroy the periodic

- 5 noise. For downsampling in text regions, a sharpening filter may be used. For upsampling of lines, a smoothing filter may be used.

#### *Downsampling Combined With Inverse Transform*

Since resampling with a linear interpolation filter is a linear operator, it can be combined with the linear wavelet transform into one operator.

- 10 Therefore, the resampling can be combined with the inverse wavelet transform into one filtering step, and resampling can be done directly on the wavelet coefficients. Since wavelet coefficients contain information on smooth and non-smooth parts of an image, different interpolation filters can be chosen depending on the absolute magnitude of wavelet coefficients or a
- 15 classifier output or on other wavelet-based criteria or without information from an external classifier.

For example, for downsampling of a one-dimensional signal with a length-4 interpolation filter  $[a_0, a_1, a_2, a_3]$  (e.g., Keys filter), given a Haar

wavelet decomposition, the interpolation filter can be implemented directly on the two finest level of the decomposition. Out of the redundant Haar wavelet coefficients, select the critically sampled coefficients with phase such that the support of the coefficients is in the length four window. In detail, in

5 order to compute an interpolation point between  $x_{2i}$  and  $x_{2i+1}$  (length-4 filter applied to window  $x_{2i-1}$  to  $x_{2i+2}$ ), the wavelet coefficients  $d_i^1, d_{i+1}^1$  at level 1 and  $d_i^2$  at level 2 and the scaling coefficient  $s_i^2$  at scale 2 have to be used. If the inverse transform step for reconstructing the sample  $x_{2i}$  is  $s_i^2 + d_i^2 + d_i^1 + d_{i+1}^1$ , then the exact interpolation filtering procedure on wavelet coefficients is

10 
$$x_{\text{int}} = (a_0 + a_1 + a_2 + a_3)s_i^2 + (a_0 + a_1 - a_2 - a_3)d_i^2 + (a_0 - a_1)d_i^1 + (a_2 - a_3)d_{i+1}^1$$

For 4-tap linear filters that preserve the DC value, i.e.  $a_0 + a_1 + a_2 + a_3 = 1$ , the above equation reduces to

$$x_{\text{int}} = s_i^2 + (a_0 + a_1 - a_2 - a_3)d_i^2 + (a_0 - a_1)d_i^1 + (a_2 - a_3)d_{i+1}^1$$

Different values for  $a_0, a_1, a_2, a_3$  yield different Keys filters. One possible  
15 way of performing a signal-adaptive downsampling in the wavelet domain is, for a given a threshold  $T$  for wavelet coefficients, select a Key filter that is

suited for text if the wavelet coefficients at level two  $|d_i^2|$  is larger than T and select a different Keys filter if  $|d_i^2| < T$ .

In one embodiment, thresholds for selecting the interpolation method or filter may vary with the amount of upsampling and for downsampling.

- 5 For example, if a particular amount of downsampling is known to be performed, an interpolation method or filter may be selected that is different than that which would be used if no such amount or a different amount of downsampling was being performed.

Similarly, the thresholds for both denoising and interpolation may

- 10 vary with the amount of upsampling and/or downsampling. The thresholds for classification and interpolation may vary with the amount of upsampling and/or downsampling. Similarly, the windows upon which they operate may vary with the amount of upsampling and/or downsampling.

- 15 In one embodiment, the filter varies with the amount of up and/or downsampling, the periodicity filtered varies with the amount of up and/or

downsampling, and/or the filter support varies with the amount of up and/or downsampling

For a one-dimensional signal of length  $N$ , choosing matrix notation, the linear interpolation transform is represented by matrix  $F_i$  and the

5 forward wavelet transform by the matrix  $W_L$ . The matrix representing the combined interpolation and the inverse transform  $W_L^{-1}$  is  $(F_i W_L^{-1})$ .

In general, given the interpolation filter and the wavelet transform, it is possible to store the filter coefficients for the combined interpolation-inverse transform in a look-up table and use that table for the filtering. If the

10 interpolation filter is of length  $N_F$  and the maximum of the length of the wavelet transform analysis lowpass and highpass filter is  $N_w$ , then the regular filter bank algorithm for the inverse transform can be used from level  $L$  down to level  $L^*$  where

$$L^* = \log(\lfloor N_{F,w} \rfloor + 1) \text{ with } N_{w,F} = L \cdot 2^{\left\lceil \left\lceil \frac{(N_w - 1)}{2} \right\rceil \right\rceil} + N_F$$

For the remaining levels  $L^*$  to 1 the modified filtering procedure represented by the matrix  $F_l^T W_L^{-1}$  has to be applied. It can be also implemented at all levels  $l < L^*$ .

A possible application for downsampling in a multifunction

- 5 machine/copier/etc. is to store the combined interpolation-inverse wavelet transform coefficients for a fixed transform and a family of interpolation filters that covers a range of sharpening and smoothing filters. If a wavelet coefficient at level  $L$  or  $L^*$  is above a given threshold (e.g., belonging to a strong text-edge), then the coefficients for a sharpening filter + inverse transform filtering is chosen. If the coefficient is below the threshold (e.g., in background regions), then the coefficients for a smoothing filter + inverse transform filtering is chosen. For the latter one it is also possible to pick one interpolation filter randomly out of the family of interpolation filter. This introduced randomness helps to destroy periodicity in areas of moiré
- 10 artifacts.
- 15

#### *Upsampling Combined With Inverse Wavelet Transform*

A similar technique can be applied for image enlargement via interpolation using a linear filter. As for resolution reduction, the length of

the interpolation filter determines the scale at which the interpolation using wavelet coefficients may be implemented. Different interpolation filters might be chosen depending on characteristics of wavelet coefficients (e.g., a smooth filter when coefficients are small, a sharpening filter when

5 coefficients are large, a filter with some random characteristics in halftone regions, etc.).

### Inverse Wavelet Transform

A normal inverse transform is performed on the processed wavelet coefficients and the scaling coefficients. In one embodiment, the inverse

10 transform is performed for all  $L$  levels. In an alternative embodiment, only  $L-L^*$  levels are inverted and the remaining  $L^*$  levels of wavelet coefficients are used for the combined inverse-downsampling transform. If  $L=L^*$ , no pure inverse is performed.

### 15 Gamma Correction

Gamma correction is not done directly on the wavelet coefficients and is performed in the traditional way in the pixel domain. Gamma correction may be performed on the scanned image, and before halftoning.

### Halftoning

After gamma correction, the halftoning process based on a wavelet-classifier may be performed. In one embodiment, dither matrices are selected depending on the classifier or changing error diffusion kernels

5 depending on the classifier.

Dithering may be used for halftoning and is commonly done by adding a pseudo random "noise" or "dither" signal to pixel intensities and then thresholding. The signal can have various property that lead to different halftoned images (Ordered Dither, Clustered Dot, Blue Noise

10 Mask).

In one embodiment, halftoning is preferred in the wavelet domain. For that purpose, a scale  $L$  is selected and White noise is added to the coefficients at scale  $l=1\dots L$ . The variance of the noise might differ from scale to scale. For example, noise variance might increase from coarse to fine

15 scales. Then an inverse transform is performed and a simple thresholding is applied to the reconstructed image to obtain a bilevel image. The noise that is added to the wavelet coefficients becomes blue noise in the pixel domain.

A prior art "blue noise mask" stores a large dither signal, for example, (256 x 256) because it is computational difficult to generate "colored" noise in

the spatial domain. However, in the wavelet domain adding noise of different intensities to different subbands provides an easy way to generate "blue noise" or any other desired "colored" noise. Thus, noise may be added in the wavelet domain and may be done so that different amounts of noise

5 are added to different bands. For example, text is usually visible in HL, LH, so less noise is added there; more noise would be added in the HH.

### An Exemplary Copier

Figure 7 is a block diagram of one embodiment of a digital copier.

10 Referring to Figure 7, digital copier 700 comprises a scanner 701 to scan a hard copy document to create a scanned image. The image processing that contains wavelet-based processing 702 is coupled to scanner 701 to receive the scanned image and generate a processed image. In one embodiment, image processing 702 comprises the processing path shown in Figure 2C. A

15 printer engine 703 is coupled to image processing to print the processed image.

In one embodiment, copier 700 includes an input to receive electronic versions of documents that may be sent directly to image processing 702 and then to print engine 703.

### An Exemplary Computer System

An image processing path may be implemented with a computer system such as shown in Figure 8. Referring to Figure 8, computer system 800 may comprise an exemplary client or server computer system in which the features described herein may be implemented. Computer system 800 comprises a communication mechanism or bus 811 for communicating information, and a processor 812 coupled with bus 811 for processing information. Processor 812 includes a microprocessor, but is not limited to a microprocessor, such as Pentium™, PowerPC™, Alpha™, etc.

System 800 further comprises a random access memory (RAM), or other dynamic storage device 804 (referred to as main memory) coupled to bus 811 for storing information and instructions to be executed by processor 812. Main memory 804 also may be used for storing temporary variables or other intermediate information during execution of instructions by processor 812.

Computer system 800 also comprises a read only memory (ROM) and/or other static storage device 806 coupled to bus 811 for storing static information and instructions for processor 812, and a data storage device

807, such as a magnetic disk or optical disk and its corresponding disk drive.

Data storage device 807 is coupled to bus 811 for storing information and instructions.

Computer system 800 may further be coupled to a display device 821,

- 5 such as a cathode ray tube (CRT) or liquid crystal display (LCD), coupled to bus 811 for displaying information to a computer user. An alphanumeric input device 822, including alphanumeric and other keys, may also be coupled to bus 811 for communicating information and command selections to processor 812. An additional user input device is cursor control 823, such
- 10 as a mouse, trackball, trackpad, stylus, or cursor direction keys, coupled to bus 811 for communicating direction information and command selections to processor 812, and for controlling cursor movement on display 821.

Another device that is coupled to bus 811 is hard copy device 824 operating as a print engine, which may be used for printing images,

- 15 instructions, data, or other information on a medium such as paper, film, or similar types of media. A scanner 830 is coupled to bus 811 to generate a scanned image.

Furthermore, a sound recording and playback device, such as a speaker and/or microphone may optionally be coupled to bus 811 for audio interfacing with computer system 800.

Note that any or all of the components of system 800 and associated hardware may be used in the present invention. However, it can be appreciated that other configurations of the computer system may include some or all of the devices.

Whereas many alterations and modifications of the present invention will no doubt become apparent to a person of ordinary skill in the art after having read the foregoing description, it is to be understood that any particular embodiment shown and described by way of illustration is in no way intended to be considered limiting. Therefore, references to details of various embodiments are not intended to limit the scope of the claims which in themselves recite only those features regarded as essential to the invention.

---

Thus, an image processing process has been described.